

Annotated Bibliography of Sources of Real-World Datasets Useful for Teaching Applied Statistics

Judith D. Singer and John B. Willett¹
Harvard University

Adkins, P. G. (1969). Deficiency in comprehension in non-native speakers. *TESOL Quarterly*, 3, 197-201.

Ninth grade student achievement (40 subjects) on tests of seventh grade material, with and without idioms and figures of speech. Useful for bivariate analysis.

Afifi, A. A., and Azen, S. P. (1979). *Statistical Analysis: A Computer Oriented Approach*, 2nd edition. New York: Academic Press.

Several extensive data sets describing the blood chemistry (cholesterol, blood pressure, etc.), cardiovascular state, socioeconomic status, and year of death. Some censored cases, could be used in the teaching of survival analysis. Other datasets include body flexibility, diet, testosterone levels in right and left testes of mice (!), weaning of rats, and infant cognitive development.

Afifi, A. A., and Clark, V. (1984). *Computer Aided Multivariate Analysis*. Belmont, CA: Lifetime Learning, 30-39.

Depression scores and selected covariates for 294 participants in the Los Angeles Depression Study. Data set includes individual item responses for a 20 question depression scale, person background characteristics and selected health variables.

Aickin, M. (1983). *Linear Statistical Analysis of Discrete Data*. New York: Wiley.

A large variety of categorical data sets including: tenure in American universities, dolphin sightings, transitions between Piagetian stages, college expectations and participation in high school athletics, political preferences, religion and marijuana, sudden infant death.

¹An earlier version of this bibliography was published in *The American Statistician*, 1990, 44(3), 223-230. The order of the authors has been determined by randomization.

Aldrich, J. H., and Nelson, F. D. (1984). *Linear Probability, Logit and Probit Models*. Sage University Paper Number 45. Beverly Hills, CA: Sage.

Data concerning the effect of the "Personalized System of Instruction" on course grades in an intermediate macroeconomics course, useful for logit analysis and log-linear modeling.

Allison, T. and Cicchetti, D. V. (1976). Sleep in mammals: Ecological and constitutional correlates. *Science*, **194**, 732-734.

Average brain weights and body weights for 62 species of mammals. Both variables are very skewed, but logarithmic transformations alleviate the skewness and improve the linearity of the scatterplot.

Andrews, D. F. and Herzberg, A. M. (1985). *Data: A Collection of Problems from Many Fields for the Student and Research Worker*. New York: Springer-Verlag.

71 sets of raw data. Many substantive areas are included, but the emphasis is generally on the physical and natural sciences. Several interesting social science examples are given, including: unemployment statistics, insurance rate information, literary data sets (the Federalist papers data set and another on Platonic prose rhythm) and the birth-day/deathday problem.

Angell, R. C. (1951). The moral integration of American cities. *American Journal of Sociology*, **53** 1-140.

Measures of the moral integration, ethnic heterogeneity, crime, welfare effort, integration and mobility of residents in 43 American cities.

American Association of University Professors (1987). The annual report on the economic status of the profession, 1986-1987. *Academe*, **73**, 1-88.

Salary data by rank, sex, and tenure status for faculty at 1,901 colleges and universities. Institutions are categorized according to Carnegie classifications.

Aylward, G. P., Harcher, R. P., Leavitt, L. A., Rao, V., Bauer, C. R., Brennan, M. J., and Gustafson, N. F. (1984). Factors affecting neobehavioral responses of preterm infants at term conceptual age. *Child Development*, **55**, 1155-1165.

Contingency table of the relationship between gestational age and neurological status for 505 babies. Also see detailed log-linear analysis of these data in: Green, J. A. (1988), Loglinear analysis of cross-classified ordinal data: Applications in developmental research, *Child Development*, **59**, 1-25.

Bain, R. (1939). Verbal stereotypes and social control. *Sociology and Social Research*, **23**, 431-446.

Multiple multiway contingency datasets on the use of proverbs and mechanical cliches used by 133 college freshmen in 1937. Data are broken down by urban/rural, gender, etc.

Barrons' Publications (1990). *Profiles of American Colleges*, Sixteenth Edition. New York: Barron's Publications.

One of many sources describing the more than 1,500 four-year colleges in this country. Relevant data include: number of applicants, number of students accepted, number of students enrolling, mean SAT scores of incoming freshman, mean class rank of incoming freshmen, faculty/student ratios, financial aid available, number of part-time students and faculty, percent of faculty with doctorates, sex composition of student body. Can be supplemented with information from American Association of University Professors salary survey and endowment data given in the *Digest of Education Statistics*.

Bell, J. C. (1914). A class experiment in arithmetic. *Journal of Educational Psychology*, 5, 467-170.

Individual data for 25 college sophomores at the University of Texas on the speed and accuracy with which they solved four types of arithmetic problems (addition, subtraction, multiplication and division).

Bell, J. C. (1916). Mental tests and college freshmen. *Journal of Educational Psychology*, 7, 381-399.

Scores on nine tests for 37 of the "best" students and 37 of the "worst" students, with notations of class rank, designed "to be of assistance to college authorities in aiding freshmen to adjust themselves to their environment" (p. 381).

Berenson, M. L., Levine, D. M., and Goldstein, M. (1983). *Intermediate Statistical Methods and Applications*. Englewood Cliffs, NJ: Prentice Hall.

A large variety of non-social science data sets (lawn service, real estate market, professional sports, foreign food), with some social science data sets scattered here and there: categorical data on health issues in children by graduating class of pediatrician, starting salaries of MBA graduates, etc.

Berndt, E. R. (1991). *The Practice of Econometrics: Classic and Contemporary*. Reading, MA: Addison-Wesley.

Dozens of data sets relevant for economics on topics such as prices of color television sets, common stocks, gold, and other commodities. An added benefit is that all the data sets are entered onto a floppy disk included with every book.

Bishop, D. W., & Ikeda, M. (1970). Status and role factors in the leisure behavior of different occupations. *Sociology and Social Research*, 54 190-208.

Multidimensional data on leisure activities by occupation. Useful for multidimensional scaling, cluster analysis, factor analysis, principal components analysis, discriminant analysis, etc.

Bock, R. D. (1975). *Multivariate Statistical Methods in Behavioral Research*. New York: McGraw Hill.

- A variety of educational growth data sets suitable for repeated measures/ MANOVA analysis, including data on responses to inkblot plates by grade and IQ over time, longitudinal (4 grades) data on scaled vocabulary scores for boys and girls, and so on. (Data repeated in Finn, J. D., & Mattsson, I. (1978). *Multivariate Analysis in Educational Research*. Chicago, IL: National Educational Resources.
- Boli, J., Allen, M. L., & Payne, A. (1985). High ability women and men in undergraduate mathematics and chemistry courses. *American Educational Research Journal*, **22**, 605-626.
- Multi-dimensional categorical data describing perceptions of course performance among men and women in physics and chemistry courses at Stanford University.
- Bullen, A. K. (1945). A cross-cultural approach to the problem of stuttering. *Child Development*, **16**, 1-88.
- Raw data for 46 children divided into four groups -- stutterers, well-adjusted, medium adjusted and poorly adjusted. Measures include age, achievement, receptivity to education, physical condition, social-personality traits, insightfulness, family background, somotype and anthropometrics.
- Chambers, J. M., Cleveland, W. S., Kleiner, B. & Tukey, P. A. (1983). *Graphical Methods for Data Analysis*. Belmont, CA: Wadsworth.
- 33 raw datasets. Many substantive areas are included, and many of these data sets are just plain interesting, such as the ages of signers of the Declaration of Independence, murder/suicides by crashing private airplanes, heights of singers in the New York Choral Society.
- Chapman, J. C. (1914). *Individual Differences in Ability and Improvement and their correlations*. Teacher's College, Columbia University Contributions to Education Number 63. New York: Teacher's College.
- Ten longitudinal data sets (10 waves) on measures of computation, color-naming and opposites-naming for 22 college-age males in New York at the turn of the century. Suitable for growth curve analysis.
- Cobb, M. C. (1917). A preliminary study of the inheritance of arithmetic abilities. *Journal of Educational Psychology*, **8**, 1-20.
- Data for on the mother, father and children in eight families, with the age of each family member and their scores on five tests (addition, subtraction, multiplication, division and copying figures). The author concludes that "it is difficult to avoid the conclusion that ... likeness is due to heredity" (p.16).
- Cooley, W. W., & Lohnes, P. R. (1985). *Multivariate Data Analysis*. Malabar, FL: Robert E. Kreiger.

Two large data sets: (1) a 20 variable subset of the PROJECT TALENT data (234 males, 271 females); and (2) the RECTANGLES data set on the dimensions of 100 rectangles, useful for factor analysis and principal components analysis.

Cox, D. R. & Snell, E. J. (1981). *Applied Statistics: Principles and Examples*. London: Chapman and Hall.

39 raw datasets. Relevant examples include: educational plans of Wisconsin school boys, statistical aspects of literary style, satisfaction with housing conditions.

Council of Great City Schools (1983). *Statistical Profiles of the Great City Schools*. Philadelphia, PA: Author.

Educational and demographic descriptors for 32 large urban school districts, including data on how these characteristics have changed over time.

Crymes, R. (1971). The relation of study about language to language performance with special reference to nominalization. *TESOL Quarterly*, 5, 217-230.

Categorical and continuous data on the number of nominals produced by native and non-native speakers. Data for experimental and control groups on pre- and post-test measures.

Devore, J. and Peck, R. (1986). *Statistics: The Exploration and Analysis of Data*. St. Paul, MN: West Publishing.

The data sets tend to be small, and many are from the sciences, but there are dozens of them. One interesting example is the movie production and promotion costs for "Dumb Movies," such as *Revenge of the Nerds* and *Police Academy*.

Draper, N. R., and Smith, H. (1981). *Applied Regression Analysis*, 2nd edition. New York: Wiley.

Some data sets with broad interest are submerged among many others, including: sex differentials in teacher pay, aptitude and age of first word, nutrition of preschoolers, ailments of university alumni, and the relationship between length of lifeline and length of life.

Dunshee, M. E. (1931). A study of factors affecting the amount and kind of food eaten by nursery school children. *Child Development*, 2, 163-183.

Data on the eating habits of 37 children, including age, sex, means (and standard deviation) for total amount of calories eaten and minutes spent at table.

Dunteman, G. H. (1984). *Introduction to Linear Models*. Beverly Hills, CA: Sage.

Data for 300 participants in the National Longitudinal Study on reading, math, gender, race, college status, SES, High school program, High school grades, creativity, stress avoidance, etc.

Eash, M. (1983). Educational research productivity of institutions of higher education. *American Educational Research Journal*, 20, 5-12.

Research productivity rankings for 25 schools of education. Measures are based on presentations at AERA and articles published in AERA journals. The paper also presents data comparing AERA program contributions in 1975 and 1976 with those in 1980 and 1981. Compare with smaller data sets presented in West, C. K. (1978). Productivity ratings of institutions based on publication in the journals of the American Educational Research Association, *Educational Researcher*, 7, 13-14 and Schubert, W. H. (1979). Contributions to AERA annual programs as an indicator of institutional productivity. *Educational Researcher*, 8, 13-17.

Educational Research Service (1987). *Scheduled Salaries for Professional Personnel in Public Schools 1986-1987*. Arlington, VA: Author.

Raw data for 1,031 school districts on enrollment, per pupil expenditures, and salaries for superintendents, central office administrators, principals, teachers, staff and support personnel.

Erickson, B. H., & Nosanchuck, T. A. (1977). *Understanding Data*. Toronto, Canada: McGraw Hill Ryerson.

An introductory textbook that melds together Tukey's exploratory data analysis and the more traditional confirmatory approaches. Many interesting data sets, including frequency of teacher criticism by student IQ, sex differences in reactions to hostile treatment by an experimenter, experimenter artifacts in social psychology research, characteristics of social networks.

Fales, E. (1933). A comparison of the vigorousness of play activities of preschool boys and girls. *Child Development*, 4, 144-157.

Age, IQ scores and activity ratings for 16 boys and 16 girls. Two activity ratings are available for each child.

Ferdinand, T. N., & Luchterhand, E. G. (1970). Inner-city youth, the police, the juvenile court, and justice. *Social Problems*, 17, 510-527.

Many multiway categorical datasets on the crimes and police disposition of inner-city youth, by race, gender, type of crime, age, etc.

Finn, J. D. (1974). *A General Model for Multivariate Analysis*. New York: Holt, Rinehart and Winston.

Raw data for four studies of relevance to the social sciences: Creativity and achievement, memory for words, essay grading practices and effects of programmed instruction.

Fleming, J. T. (1967). The measurement of children's perception of difficulty in reading materials. *Research in the Teaching of English*, 1, 136-156.

Individual data on 60 subjects. Data include background characteristics (gender, age, IQ, etc.) and achievement (comprehension, reading, word knowledge, spelling, language,

study skills) for school-age children. Suitable for multiple regression, multivariate analysis, etc.

Fox, J. (1984). *Linear Statistical Models and Related Methods with Applications to Social Research*. New York: John Wiley.

Several interesting data sets including: relationship between status, authoritarianism and conformity, methods to enhance recall of words, causes of the 1907 Romanian Peasant rebellion.

Fraumeni, J. F. Jr. (1968). Cigarette smoking and cancers of the urinary tract: Geographic variation in the United States. *Journal of the National Cancer Institute*, **41**, 1205-1211.

Aggregate cigarette smoking and cancer death rates, by type of cancer and state.

Garwood, A. N. (Ed.) (1986). *Massachusetts Municipal Profiles*. Wellesley Hills, MA: Information Publications.

Sociodemographic characteristics for 353 Massachusetts towns and cities, including data on age, race, sex, income, labor force participation, voter registration, police, fire, crime, taxation, libraries and schools. The company publishes similar books for other states; write to them at Box 356, Wellesley Hills, MA 02181.

Gastwirth, J. L. (1984). Statistical methods for analyzing claims of employment discrimination. *Industrial and Labor Relations Review*, **38**, 75-86.

Several data sets used in employee discrimination cases against organizations including the University of Texas at Dallas, the city of Wichita Falls, and the Federal Reserve Bank.

Gelb, S. A., & Mizokawa, D. T. (1986). Special education and social structure: The commonality of "exceptionality." *American Educational Research Journal*, **23**, 543-557.

State-level data on percentage of children in each category of special education and sociodemographic composition of the states. Washington, DC is a high-leverage outlier for the relationship between percent of students classified as educably mentally retarded and percent of population that is black.

Gerlach, M. (1939). A study of the relationship between psychometric patterns and personality types. *Child Development*, **10**, 269-278.

Raw data for 61 maladjusted children on two IQ tests, as well as information on their sex, age, parentage (both Foreign, both American, Mixed) and maladjustment type (aggressive or asocial). Authors explore relationship between IQ and all these predictors.

Gnanadesikan, R. (1977). *Methods for Statistical Data Analysis of Multivariate Observations*. New York: John Wiley.

Several multivariate data sets from a variety of disciplines including engineering, manufacturing, biology and mining. The volume includes several well-known data sets such as Fisher's iris data (1936) and Rothkopf's Morse-code confusion data (1957) which

have utility for the teaching of principal components analysis, factor analysis, multidimensional scaling and cluster analysis.

Gordon, N. J., Nucci, L. P., West, C. K., Hoerr, W. A., Vguroglu, M., Vukosavich, P., & Tsai, S. L. (1984). Productivity and citations of educational research: Using educational psychology as the data base. *Educational Researcher*, **13**, 14-20.

Citation frequencies and dates of birth for 187 prominent educational researchers.

Gulliksen, H. (1934). A rationale equation of the learning curve, based on Thorndike's law of effect. *Journal of General Psychology*, **11**, 395-434.

Longitudinal data on rats -- some brighter than others -- trained to discriminate 18 cm and 12 cm circles.

Haberman, S. J. (1978). *Analysis of Qualitative Data*. New York: Academic Press.

Several categorical data sets of wide interest: Suicides by day of the week, homicides by month, stressful events, etc.

Hahn, H. H., & Thorndike, E. L. (1914). Some results of practice in addition under school conditions. *Journal of Educational Psychology*, **5**, 65-83.

Individual data from an experiment on the effects of time lapsed between pre-tests and post-tests for 167 students in grades 4, 5, 6, and 7.

Hamilton, L. C. (1990). *Modern Data Analysis: A First Course in Applied Statistics*. Pacific Grove, CA: Brooks-Cole.

Dozens of interesting data sets on topics such as the unemployment of graduates from 44 British universities, the drinking habits of 120 mothers and their adolescent children, average salaries and overall performance of 26 baseball teams, and marriage rates in the 50 states (Nevada is a very high outside value).

Hand, D. J. & Taylor, C. C. (1987). *Multivariate Analysis of Variance and Repeated Measures: A Practical Approach for Behavioral Scientists*. London: Chapman and Hall.

Raw data from 8 studies in psychology and psychiatry, on topics as diverse as headaches, smoking and Alzheimer's disease.

Harris, J. A., Jackson, C. M., Paterson, D. G., & Scammon, R. E. (1930). *The Measurement of Man*. Minneapolis, MN: University of Minnesota Press.

Many unique and interesting data sets on the link between physical and psychological characteristics. Do blondes have more fun? Do lunatics eyebrows join together in the middle? Are manic depressives thin or fat?

Harris, S., & Harris, L. B. (1986). *The Teacher's Almanac*. New York: Facts on File.

Assorted education data by state and school district, including teacher salaries, high school graduation rates, functional illiteracy rates, and presence of computers in schools.

- Heyneman, S. P. (1976). Influences on academic achievement: A comparison of results from Uganda and more industrialized societies. *Sociology of Education*, **49**, 200-211.
- Aggregate data on school achievement and economic development for 18 nations. Includes measures of preschool influence, GNP, percent enrollment in primary and secondary school.
- Hirsch, N. D. M. (1928). An experimental study of the East Kentucky mountaineers: A study in heredity and environment. *Genetic Psychology Monographs*, **3**, 183-244.
- Ages and IQ scores of siblings in 44 families.
- Hirsch, N. D. M. (1930). An experimental study upon three hundred school children over a six-year period. *Genetic Psychology Monographs*, **7**, 487-546.
- IQ scores over a six year period for 343 children.
- Hodges, J. L., Krech, D., & Crutchfield, R. S. (1975). *Statlab: An Empirical Introduction to Statistics*. New York: McGraw-Hill.
- Data on two parents and one child from each of 1296 families. Variables include information on height, weight, and psychological tests.
- Hollander, M., & Proschan, F. (1984). *The Statistical Exorcist: Dispelling Statistics Anxiety*. New York: Marcel Dekker.
- A host of data sets of different sizes on many different topics including: blood pressure and obesity of Mexican-Americans, baseball data, 1970/1 draft lottery, promotion rates among male and female pharmacists, leisure time companions of black women, ranking of rum brands by different nationalities, preference for Charlie's Angel's actors, longevity and environment, color of canned tuna, etc.
- Howard, G. S., Cole, D. A., & Maxwell, S. E. (1987). Research productivity in psychology based on publication in the journals of the American Psychological Association. *American Psychologist*, **42**, 975-986.
- Productivity, reputation and size of psychology departments at 75 universities.
- Izenman, A. J. (1972). Reduced rank regression for the multivariate linear model. Doctoral dissertation, Department of Statistics, University of California at Berkeley.
- Body length of crickets as a function of geographical location and weather throughout the USA.
- Jensen, A. R. (1974). Kinship correlations reported by Sir Cyril Burt. *Behavioral Genetics*, **4**, 1-28.
- "Burt's final assessments" of IQs of monozygotic twins reared apart, with "social class" ratings of the homes. For information on Burt's falsification of the data, see Dorfman, D. D. (1978). The Cyril Burt Question: New Findings. *Science*, **201**, 1177-1186; other sources include correspondence related to Dorfman's article: Stigler, S. M. (1979).

- Letter to the editor. *Science*, **204**, 242-245; Rubin, D. B. (1979). Letter to the editor. *Science*, **204**, 245-246; Dorfman, D. D. (1979). Letter to the editor. *Science*, **204**, 246-254; Hearnshaw, L. S. (1979). *Cyril Burt: Psychologist*. London: Hodder and Stoughton, Ch. 12.
- Jensen, A. R. (1970). IQ's of identical twins reared apart. *Behavioral Genetics*, **1**, 133-148.
Original data from four studies of IQs of identical twins reared apart: Newman, Freeman & Holzinger (1937), Shields (1962), Juel-Neilsen (1965) and Burt (1955).
- Johnson, B., & Courtney, D. M. (1931). Tower building. *Child Development*, **3**, 161-162.
Twenty-five children were asked to build towers on each of two occasions. Each time they were given: (a) a set of cubes; and (b) a set of cylinders. Raw data are given on the number of blocks of each type used each time, and how many minutes it took to construct the tower.
- Johnson, W. D., & Koch, G. G. (1971). A note on the weighted least squares analysis of the Ries-Smith contingency table data. *Technometrics*, **13**, 438-447.
Large four-way contingency table derived from a randomized experiment in which several brands of detergent were compared under a variety of user-related conditions.
- Jones, L. V., Lindzey, G., & Coggeshall, P. E. (1982). *An assessment of research-doctorate programs in the United States: Social Sciences*. Washington, DC: National Academy Press.
"Quality" rankings and characteristics of university departments in the social sciences, by discipline. Data include number of faculty, number of students, productivity of faculty, number of grants awarded, follow-up placement of doctoral students.
- Karelitz, S., Fisichelli, V. R., Costa, J., Karelitz, R., & Rosenfeld, L. (1964). Relation of crying in early infancy to speech and intellectual development at age three years. *Child Development*, **35**, 769-777.
Data for 38 infants on crying in early infancy and later measures of IQ.
- Koch, H. L. (1933). Popularity in preschool children: Some related factors and a technique for its measurement. *Child Development*, **4**, 164-175.
Popularity scores for 17 children: percent each child was named first, percent each child name last, effects of ordering and the effects of sex.
- Kroc, R. J. (1984). Using citation analysis to assess scholarly productivity. *Educational Researcher*, **13**, 17-22.
Citation frequency data for 51 schools of education. Measures include mean citation rate, percentage of faculty with 10-100 citations and percentage of faculty with no citations. Kroc correlates these measures with five rankings of schools of education; these analyses could be recreated by abstracting data from the five sets of rankings cited in his bibliography.

- Leinhardt, G., & Leinhardt, S. (1980). Exploratory data analysis: New Tools for the analysis of empirical data. *Review of Research in Education*, **8**, 85-157.
- Three measures of reading instruction for a sample of 53 learning disabled students, by curricular approach and school.
- Leinhardt, S. and Wasserman, S. S. (1979). Teaching regression: An exploratory approach. *The American Statistician*, **33**(4), 196-203.
- Life expectancy and per capita income for 105 nations divided into five wealth classifications (industrialized, petroleum exporting, higher, middle and lower).
- Louano-Kerr, J., Semles, V., & Zimmerman, E. (1977). A profile of art educators in higher education: Male/female comparative data. *Studies in Art Education*, **18**, 21-37.
- Multiway contingency table data on the qualifications (Masters vs. doctorate) of art educators, by year and gender.
- Maresh, M. M., & Deming, J. (1939). The growth of the leg bones in 80 infants: Roentgenograms versus anthropometry. *Child Development*, **10**, 91-106.
- Individual data for 80 children on the sizes of 10 bones, measured by both x-rays and anthropometry at each of 3, 4, or 5 occasions, by sex. The authors construct lots of individual growth curves.
- Mason, T. J., & McKay, F. W. (1974). US Cancer Mortality by County: 1950-1969. Washington, DC: US Government Printing Office.
- Lung cancer mortality by degree of urbanization and gender, in Louisiana.
- Mickey, M. R., Dunn, O. J., & Clark, V. (1967). Note on the use of stepwise regression in detecting outliers. *Computers and Biomedical Research*, **1**, 105-109.
- Gesell adaptive scores and age at first word (in months) for 21 children with cyanotic heart disease. The data set contains some interesting outliers and high leverage cases.
- Mosteller, F., Hyman, H., McCarthy, P. J., Marks, E. S., & Truman, D. B. (1949). *The Pre-Election Polls of 1948*. New York: Social Science Research Council.
- "Dewey beats Truman" is one of the classic headlines of the twentieth century. The data contained in this book provide an opportunity to examine the pre-election polls to ask why this headline seemed plausible. Data are provided by the respondents' social class, as well as the month in which the pre-election poll was conducted. Also see analyses in Baker, S. G., & Laird, N. M. (1988), Regression analysis for categorical variables with outcome subject to nonignorable nonresponse, *Journal of the American Statistical Association*, **83**, 62-69.
- Mosteller, F. & Tukey, J. W. (1977). *Data Analysis and Regression: A Second Course in Statistics*. Reading, MA: Addison-Wesley.

Raw data for 13 data sets across several disciplines. Relevant examples include a subset of 20 from the Coleman Report, educational expenditures for Massachusetts school districts, municipal bond data for 20 US cities.

National Center for Education Statistics (1987). *The Condition of Education*. Washington, DC: US Department of Education, and National Center for Education Statistics (1987). *Digest of Education Statistics*. Washington, DC: US Department of Education.

Annual reports issued by the Department of Education providing descriptive information on education, often over time, sometimes by state, occasionally by school district. The data on university endowments can be used in conjunction with other university level data, such as that given in Barron's (1987).

National Education Association (1965). Annual report on public-school financing. *NEA Research Bulletin*, **43**, 90-92.

State-by-state listing of per-pupil expenditures (in 1955-56 and 1964-65) and percent of school revenues from state and local sources. The entire journal series contains a wealth of information on class size, salaries, public expenditures for education, etc.

National Education Association (1926). The ability of the states to support education. *Research Bulletin of the National Education Association*, **4**.

Several chapters of state-level data documenting wealth, per capita income, educational expenditure, population size, number of adults per child, and so on.

National School Boards Association (1986). *A Survey of Public Education in the Nation's Urban School Districts*. Alexandria, VA: author.

Data for 61 school districts on educational policies and practices, as well as selected education and economic descriptors.

Nelson, W. (1982). *Applied Life Data Analysis*. New York: John Wiley.

Several data sets presenting time-to-failure of products, useful for examining industrial product reliability.

Neuman, S., & Ziderman, A. (1985). Do universities maintain common standards in awarding first degrees with distinction?: The case of Israel. *Higher Education*, **14**, 447-459.

Number of first degree graduates with and without distinction, by school and faculty within school. Suitable for log-linear modeling.

Opening lines (1985, September). *Harper's*, pp. 29-30.

Selected results from a study of opening lines used in singles bars in the St. Louis area. Two-way contingency table describing the relationship between type of opening line (compliments, propositions, etc.) and time of evening.

Phillips, D. P. (1978). Deathday and birthday: an unexpected connection. In Tanur, J.M., Mosteller, F., Kruskal, W. H., Link, R. F., Pieters, R. S., Rising, G. R., and

Lehmann, E. L. (eds.), *Statistics: A Guide to the Unknown*, 2nd ed. San Francisco, CA: Holden-Day, 71-85. See, also, Phillips, D. P. (1977), Motor vehicle fatalities increase just after publicized suicide stories, *Science*, **196**, 1464-1465; Phillips, D. P. (1978), Airplane accident fatalities increase just after newspaper stories about murder and suicide, *Science*, **201**, 748-750; Phillips, D. P., and Carstensen (1986), Clustering of teenage suicides after television news stories about suicide, *New England Journal of Medicine*, **315**, 685-689 (and related articles in this issue); Schultz, R., Bazerman, M. (1980), Ceremonial occasions and mortality: A second look, *American Psychologist*, **35**, 253-261.

David Phillips has made a cottage industry of looking at what many might term coincidences -- birthdays and deathdays and copycat suicides after popularized accounts in the media. These are but a handful of articles, each listing the detailed raw data on deaths following these events that led him to his conclusions.

Plackett, R. L. (1981). *The Analysis of Categorical Data*. New York: MacMillan).

A large variety of categorical data sets including: fingerprints, family size, work conditions and work quality, behavioral problems and birth order, high school rank by gender and socioeconomic status.

Powell, B. & Steelman, L. C. (1984). Variations in state SAT performance: Meaningful or misleading? *Harvard Educational Review*, **54**, 389-412.

Mean SAT scores and percent of high school seniors taking the SAT, by state for 1982. For additional data, and a critique of their analyses, see: Wainer, H, Holland, P. W., Swinton, S., & Wang, M. H. (1985), On "State Education Statistics", *Journal of Educational Statistics*, **10**, 293-325. Also see: Rosenbaum, P. R. & Rubin, D. B. (1985), Discussion of "On State Education Statistics": A difficulty with regression analyses of regional test score averages, *Journal of Educational Statistics*, **10**, 326-333. and Wainer, H. (1986), Five pitfalls encountered while trying to compare states on their SAT scores, *Journal of Educational Measurement*, **23**, 69-81.

Rubin, E. (1972). Statistical exploration of a medieval household book, *The American Statistician*, **26**, 37-39.

Number of meals served, breads baked and ale brewed at the de Bryene household from October 1412-September 1413, by month. That's right, the fifteenth century.

Ryan, B. F., Joiner, B. L., & Ryan, T. A. Jr. (1985). *Minitab Manual*, 2nd edition. Boston, MA: Duxbury.

Thirty data sets of small to moderate sizes, on topics ranging from education to cartoons. The educational data sets include information on school strikes and freshman SAT verbal and math scores.

Sandvern, J. (1971). Causes of lacking sense of well-being in school. *Scandinavian Journal of Educational Research*, **15**, 21-60.

- Many multiway contingency tables dealing with students' liking of school in Norway, broken down by type of school, SES, etc.
- Scarcella, R. C. (1984). How writers orient their readers in expository essays: A comparative study of native and non-native english writers. *TESOL Quarterly*, 671-688.
- Categorical data on the language background and language proficiency of native and non-native speakers and how this influences their choice of writing device.
- Shearer, L. (1987). How will history rate Nancy Reagan? *Parade Magazine*, 14 June 1987, p. 8.
- Rankings for 17 first ladies from Florence Harding through Nancy Reagan on 10 dimensions ranging from integrity, leadership and accomplishments.
- Skodak, M., & Skeels, H. M. (1949). A final follow-up study of one hundred adopted children. *Journal of Genetic Psychology*, 75, 85-125.
- Raw data for 100 children who were adopted at birth. Measures include: natural mother's IQ and education level, foster mother's IQ and education level, foster father's occupation and child's IQ on each of 5 occasions, from infancy through pre-adolescence.
- Stevens, J. (1986). *Applied Multivariate Statistics for the Social Sciences*. Hillsdale, NJ: Lawrence Erlbaum.
- Many artificial data sets as well as approximately 20 interesting small to moderately sized real data sets, including: pre/post data on the influence of Sesame Street, risk of reading problems among kindergartners, behavior reversal, programmed music instruction of elementary school children, IQ testing, etc.
- Stewart, L. H. (1955). The expression of personality in drawings and paintings. *Genetic Psychology Monographs*, 51, 45-103.
- Data for 28 boys given including: 2 IQ scores (Terman and Stanford-Binet), 2 socioeconomic status measures (parents education and father's occupation), somotypes (endomorph, mesomorph, ectomorph), and drawing type.
- Supreme court ruling on death penalty (1988). *Chance*, 1, 7-8.
- Three-way contingency table on the relationship between race of victim, race of defendant and use of the death penalty, showing that the death penalty is not uniformly applied.
- Taagepera, R. (1972). The size of national assemblies. *Social Science Research*, 1, 385-401.
- Aggregate data on the size of the national assemblies in more than 100 countries, with data on population, literacy, working age, etc.
- Tanner, M. A. (1990). *Investigations for a Course in Statistics*. New York: MacMillan.

Dozens of exercises for generating raw data from students in the class. on variables such as the distance between eyebrows, and the ability to judge the lengths of lines and distances between points.

Timm, N. H. (1975). *Multivariate Analysis with Applications in Education and Psychology*. Belmont, CA: Wadsworth.

Raw data for a handful of data sets gathered in educational settings, including: effects of delay in oral practice on second language learning (pp. 228-229), relationship between recall and sentence structure (p. 233), predictors of student performance on the Peabody Picture Vocabulary Test (p. 281).

Tufte, E. R. (1974). *Data Analysis for Politics and Policy*. Englewood Cliffs, NJ: Prentice-Hall.

Several real data sets on topics of political and social relevance, thoroughly analyzed using fairly elementary statistical methods. Topics include the relationship between bureaucracy size and population, the number of radio receiver licenses issued and the proportion of mental defectives in Britain (per 100,000 population), and the relationship between dietary fat consumption and death rates in selected countries around the world.

Tufte, E. R. (1978). Registration and Voting. In Tanur, J. M. et al., *Statistics: A Guide to the Unknown*, 2nd edition. San Francisco: Holden-Day), 195-204.

Data on percent on population registered and percent of population voting in the 1960 election for 104 cities. Data are analyzed in greater detail in: Kelly, S. Jr., Ayres, R. E. and Bowen, W. G. (1967). Registration and Voting: Putting first things first. *American Political Science Review*. **61**, 359-379.

United Nation's Children's Fund (1987). *The State of the World's Children*. New York: Oxford University Press.

Sociodemographic, education, health and economic indicators for 130 countries.

Walberg, H. J., & Rasher, S. P. (1974). Public school effectiveness and equality: new evidence and its implications. *Phi Delta Kappan*, **66**, 3-9, and Walberg, H. J., & Rasher, S. P. (1976). Improving regression models. *Journal of Educational Statistics*, **1**, 253-277.

The authors analyze data for the 50 states on the relationship between failure on the selective service exam administered during 1969-1970 and contextual and education descriptors of the states. The selection bias inherent in analyses of state level SAT scores is present here, but the data set is interesting.

Weisberg, S. (1980). *Applied Linear Regression*. New York: Wiley.

Raw data for several interesting data sets including Cyril Burt's IQ data, Allison and Cicchetti's brain weight and body weight data, three time points for 26 boys and 32 girls who participated in the Berkeley Guidance study (anthropometric information only, however.)

Whiting, J. M. & Child, I. L. (1962). *Child Training and Personality*. New Haven, CN: Yale University Press.

Many characteristics in dozens of societies around the world, including age at weaning, toilet training, fear of ghosts, rituals, etc.

Wilson, M. E., & Mather, L. E. (1974). Life expectancy [Letter to the editor]. *Journal of the American Medical Association*, **229**(11) 1421-1422.

Age of person at death (in years) and the length of the person's lifeline (in centimeters) for 50 individuals. Not surprisingly, the test of $H_0: r=0$ cannot be rejected.

Zeligs, R. (1938). Tracing racial attitudes through adolescence. *Sociology and Social Research*, **23**, 45-54.

Chronological, mental age, social distance scores, and gender for twelve children tested three times during adolescence. Also indices for racial tolerance over three occasions of measurement for 39 racial groups.

Zimmerman, J. (1917). The Binet-Simon Scale and Yerkes Point Scale: A comparative examination of 100 cases. *Journal of Educational Psychology*, **8**, 551-558.

Individual data for 100 students on these two IQ tests, with information on student sex, age and native language.

Judith D. Singer and John B. Willett are Associate Professors at the Harvard University Graduate School of Education specializing in quantitative methods. They have written and presented dozens of talks, workshops, and papers on the application of statistical methods in education and the social sciences. Together with other colleagues, they have written two books, *By Design?* and *Who Will Teach?*, both published by Harvard University Press. Their current research interests center on applications of survival analysis in the social sciences. They have recently received the Raymond B. Cattell Award and the Palmer O. Johnson Award from the American Educational Research Association and an NSF Visiting Fellowship from the ASA.